Theoretical Studies on the Thermodynamic Product Size Distribution in Nucleation−Elongation Polymerization under Imbalanced Stoichiometry

Dahui Zhao* and Kan Yue

Beijing National Laboratory for Molecular Sciences (BNLMS), Department of Applied Chemistry, and the Key Laboratory of Polymer Chemistry and Physics of the Ministry of Education, College of Chemistry and Molecular Engineering, Peking University, Beijing 100871, China

Received August 21, 2007; Revised Manuscript Received February 26, 2008

ABSTRACT: Three different models are developed to calculate the thermodynamic product size distribution in a nucleation−elongation polymerization between a pair of A−A and B−B typed comonomers. These monomers are designed to undergo a single step of nucleation prior to an isodesmic chain elongation, namely, a cooperative, step-growth polymerization with dimerization being an energetically less favored process. Particularly, emphasis is laid on analyzing product distribution under conditions of imbalanced functionality stoichiometry. Consistent results are obtained from independent approaches, mechanistic and statistical, demonstrating that when the mole ratio of the comonomers deviates from unity, at polymerization equilibrium such a nucleation−elongation polymerization generates products of substantially higher molecular weights than those from a corresponding isodesmic system having an identical energetics for chain propagation yet without the nucleation process. This higher molecular weight is shown achieved by retaining a large portion of the excess monomer unreacted at equilibrium and selectively compose product chains with comonomers at a roughly stoichiometric ratio. Essentially, such a polymer−monomer coexisting bimodal distribution is a result from destabilization of the oligomeric species due to the nucleation effect.

Introduction

Nucleation−elongation (N−E) is a ubiquitous polymerization mechanism in the biological regime, typically associated with helical or tubular structures, and has been relatively well studied, e.g., with actin, tubulin, bacterial flagellin, and tobacco mosaic virus.[1,2] This N−E mechanism assumes a cooperative chain growth, featuring a slow, energetically less favored initiation process forming a nucleus followed by an accelerated, energy-favorable chain elongation. More recently, chemists have achieved a diversity of supramolecular complexes via the assembly of synthetic building blocks emulating the helical or tubular architectures exhibited by the natural entities,[3] and some of them have even proven to display a similar nucleated polymerization mechanism as manifested by their biological counterparts.[3e] Both theoretical and experimental investigations have shown that a polymerization following the N−E mechanism exhibits distinct thermodynamic and kinetic characteristics from those of isodesmic systems.[1,4] It is thus envisioned that systems governed by such N−E mechanism will display novel properties that are not accessible with classical synthetic polymerization techniques.[3e,4–6] Nonetheless, most previous theoretical studies aimed at elucidating thermodynamic and/or kinetic properties of N−E polymerizations have focused on systems involving a single A−B (or A−A) type, self-complementary monomer.[1,7] Apparently, in these systems the stoichiometry of reacting functional groups is intrinsically balanced. In great contrast to the extensive studies and discussion on such self-complementary monomer systems, theoretical studies devoted to N−E polymerization of a pair of comonomers that each bear two functionalities of the same kind and would only react with one another (i.e., A−A and B−B type comonomers) have so far been limited.[4,7] Naturally, investigations at the effects of stoichiometry imbalance on the N−E polymerization is even scarce.

The motivation for such a theoretical inspection stems from our previous experimental endeavors at creating a synthetic, helical architecture by polymerizing a pair of *m*-phenyleneethynylene comonomers via an imine metathesis condensation reaction.[6,8] In this system, a N−E chain growth is believed to arise from the product chain folding into a helical structure. Two notable features emerged from this reaction, representing typical differences between synthetic and biological polymerizations. First, unlike protein polymers comprising subunits joined by supramolecular interactions, monomers in this synthetic system were covalently connected via a condensation reaction, and a small-molecule condensation byproduct was generated during the polymerization. The general consequences of the presence of such a small-molecule byproduct in the N−E polymerization have been discussed in an earlier contribution.[4] Second, a pair of A−A and B−B typed comonomers were employed in this synthetic polymerization. One of the unique opportunities that were opened up by harnessing such an A−A and B−B comonomer pair is that the N−E polymerization may now be conducted under conditions of imbalanced stoichiometry. This subject was of particular interest to us because we predicted that a polymerization experiencing a nucleation event during chain growth should be more tolerant of stoichiometric imbalance of the reacting functional groups. Namely, polymers of substantial molecular weights (MW) should be generated from a reversible N−E polymerization even when the functional group ratio is not rigorously controlled to match the reaction stoichiometry.[8,9] This expectation is in great contrast to that for a classical isodesmic, step-growth polymerization, wherein the product MW is extremely sensitive to functionality stoichiometry balancing and high polymers are virtually unattainable with a significantly imbalanced stoichiometry.[10] This is because, in a nucleated polymerization, species of similar sizes as nuclei are of higher free energy and less stable relative to longer chains and monomers. Thus, a group of short oligomers is energetically disfavored compared to longer polymers that coexist with unreacted monomers. In the equilibration of a N−E polymer-

---

* Corresponding author. E-mail: dhzhao@pku.edu.cn.

10.1021/ma071884a CCC: $40.75

ization, the product distribution would thus be shifted to favor more stable species and bifurcate into polymers and monomers at the expense of depleting unstable oligomers, giving rise to a bimodal distribution.

This is indeed what was observed with a condensation polymerization that exhibits the N−E chain growth under conditions of imbalanced stoichiometry. Experimental results from this *m*-phenyleneethynylene polymerization supported our hypothesis by generating polymer products with substantial MW at highly imbalanced functional group stoichiometries and leaving a significant amount of monomers unreacted at equilibrium.[8] A simple model capturing the fundamental mechanistic features of this nucleated polycondensation system also gave supportive evidence for the powerful capacity of the N−E polymerization in generating polymers under imbalanced stoichiometry. However, because of its oversimplified formulation, information on polymer product size distribution was not available from this previous model.[8] In the current report, more rigorous approaches are taken to calculate the product MWs and their distribution under representative conditions of N−E polymerization with an imbalanced stoichiometry. Results from these studies have supported our original hypothesis by illustrating that the N−E polymerization featuring a single nucleation step (i.e., dimerization) in chain growth can yield polymeric products of significantly higher MWs than those in an isodesmic polymerization when the two systems have an identical driving force for chain propagation beyond dimerization.

Two methods for analyzing product size (or MW) distribution in step-growth polymerizations are explored in the present study—mechanistic[1,4,7a–e] and statistical approaches.[7f,10] In designing new models to study the properties of N−E systems, comonomers bearing distinct functional groups must be differentiated. A mechanistic model depicting a polymerization between a pair of A−A and B−B monomers exhibiting a single nucleation step will first be elaborated to reveal the thermodynamic size distribution of products at varied comonomer stoichiometries. Then, statistical methods, including the recursive approach developed by Macosko and Miller,[10b] will be applied to the same system. Agreement will be shown reached by these different approaches. In order to explicitly delineate the effect of nucleation, for both mechanistic and statistical models comparisons will be made between the isodesmic and N−E mechanisms.

It should be noted that, although the following discussion will be focused on calculation results concerning polymerizations displaying a nucleating dimerization process, mathematical models developed herein is equally applicable to systems that exhibit a tight (i.e., energetically favored) dimerization. Additionally, simple modifications to the models will avail corresponding results for polymerizations that do not generate the small-molecule byproduct. Relevant results are shown in the Supporting Information.

Results and Discussion

I. Mechanistic Model. The mechanistic scheme shown by eq 1 describes a reversible step-growth polymerization as a process of sequential monomer additions to a growing species, leading to a set of products of different sizes. A and B represent a pair of bisfunctional comonomers, each bearing two functional groups of the same kind. A_mB_n represents a species comprising *m* A units and *n* B units in an alternating sequence, since A and B only react with each other but not with themselves (i.e., $m = n$ or $n \pm 1$). D is a small-molecule byproduct that is generated with each condensation reaction. Each monomer addition step in eq 1 is reversible and associated with an equilibrium constant. Under equilibrated conditions, the concentration of each product sequence, denoted by species in brackets, $[A_mB_n]$, can thus be expressed as a function of concentrations of unreacted monomers A and B and byproduct D, along with a set of association constants (eq 2). Apparently, the formation of a given species can be realized via a number of different pathways, but its equilibrium concentration is independent of these specific pathways. The particular set of equations chosen in eq 1 is not necessarily the kinetically most favorable pathways but representative, thermodynamically viable ones.

$$\begin{gathered} A + B \overset{4\sigma K}{\rightleftharpoons} AB + D \\ AB + A \overset{K}{\rightleftharpoons} A_2B + D \\ AB + B \overset{K}{\rightleftharpoons} AB_2 + D \\ A_2B + B \overset{4K}{\rightleftharpoons} A_2B_2 + D \\ \vdots \end{gathered} \tag{1}$$

$$\begin{gathered} [AB] = 4\sigma K[A][B]/[D] \\ [A_2B] = K[AB][A]/[D] = 4\sigma K^2[A]^2[B]/[D]^2 \\ [AB_2] = K[AB][B]/[D] = 4\sigma K^2[A][B]^2/[D]^2 \\ [A_2B_2] = 4K[A_2B][B]/[D] = 4^2\sigma K^3[A]^2[B]^2/[D]^3 \\ \vdots \\ [A_nB_n] = \sigma 4^n K^{(2n-1)}[A]^n[B]^n/[D]^{(2n-1)} \\ [A_nB_{n+1}] = \sigma 4^n K^{2n}[A]^n[B]^{n+1}/[D]^{2n} \quad (n \geq 1) \\ [A_{n+1}B_n] = \sigma 4^n K^{2n}[A]^{n+1}[B]^n/[D]^{2n} \end{gathered} \tag{2}$$

In a realistic polymerization system (including the N−E polymerization), the association constants for monomer additions to growing chain ends of varied sizes may or may not be the same. In the current model described by eq 1, in order to simplify the calculations and elucidate the fundamental features of the N−E polymerization, only the association between monomers A and B (i.e., dimerization) is set to have a free energy change different from those of subsequent chain elongation steps. The magnitude of this disparity in free energy change is reflected as a coefficient, σ, in the equilibrium constant of dimerization. All further monomer additions to dimer and higher species are assumed to take place with an identical free energy change and thus have the same equilibrium constant, *K*. Evidently, when $\sigma < 1$ eq 1 corresponds to a N−E polymerization, and if $\sigma = 1$ it represents a classical isodesmic step-growth system. Moreover, as implied by eq 1, A and B are designated to be energetically equivalent.

It should be noted that a cofactor 4 is added in the association constants of some of the steps in eq 1. This is a result of statistical factors in the polymerization equilibration. If the reaction between a functionality A (functional groups in monomer A will be referred to as functionality A in the following discussion; likewise for functionality B in monomer B) and a functionality B assumes an equilibrium constant *K*, then the association of a bifunctional monomer A with a bifunctinoal monomer B has a statistical factor of 4 because each of the two functionality A may react with either of the two functionality B. Such a statistical factor should apply to monomer additions to all symmetrical sequences (i.e., A_nB_{n+1} and $A_{n+1}B_n$, each having the same kind of end functionalities). For monomer additions to asymmetrical species, A_nB_n, a

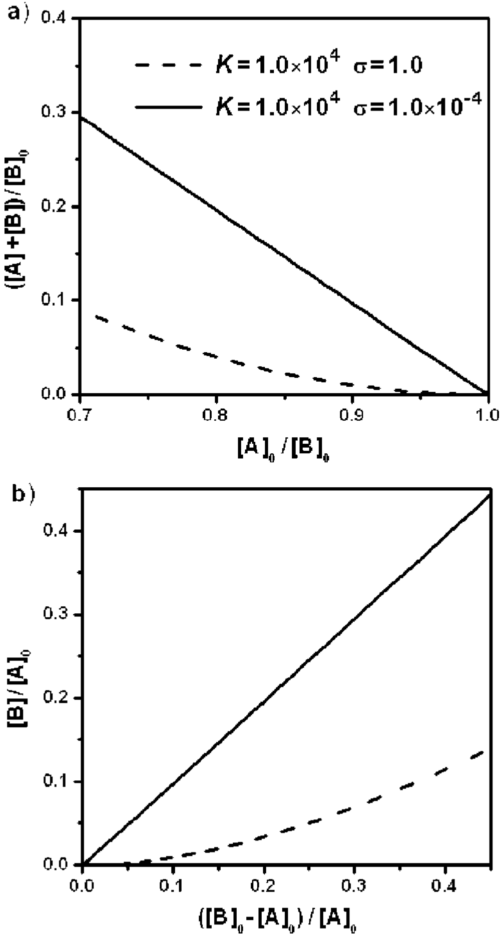


Figure 1. Plots of the relative concentration of unreacted monomer(s) at polymerization equilibrium as a function of (a) the initial monomer molar ratio and (b) the relative amount of comonomer in excess in the nucleation−elongation (solid line) and isodesmic (dashed line) polymerizations, respectively; representative values of equilibrium constant K and nucleation factor σ used for calculations are as indicated (the denotations of K and σ are defined in eq 1).

statistical factor of 2 for the forward reaction is canceled out by the same statistical factor for the reverse reaction (due to the fact that two equivalent bonds in the product may break up to give back the same set of starting materials).

On the basis of eq 2 and the total mass equations (eqs S3−S5 in the Supporting Information), one may calculate [A], [B], and [D] at polymerization equilibrium as functions of the total (initial) concentrations of A and B (denoted by $[A]_0$ and $[B]_0$, respectively) as well as σ and K.[11] As an equal-molar reaction, $[A]/[A]_0$ and $[B]/[B]_0$ are independent of absolute values of $[A]_0$ and $[B]_0$ and only dependent on $[A]_0/[B]_0$ at given σ and K values. In order to demonstrate the effect of monomer stoichiometry change on the polymerization, correlation between the concentration of unreacted monomers at equilibrium, ([A] + [B]), and the initial monomer ratio, $[A]_0/[B]_0$, was studied, and representative results are shown in Figure 1. Under representative conditions corresponding to isodesmic and N−E mechanisms ($\sigma = 1.0$ and 1.0×10^{-4} for the isodesmic and N−E systems, respectively, and $K = 1.0 \times 10^4$ for both), distinct thermodynamic behaviors are manifested.[12] First of all, within the examined range of $[A]_0/[B]_0$ (from 1.0 to 0.7), the relative concentration of unreacted monomers, $([A] + [B])/[B]_0$, at equilibrium of the N−E polymerization was consistently found to be much higher than that in the corresponding isodesmic system (Figure 1a).[13] This result explicitly indicates that, even though the two mechanisms have an infinite number of steps proceeding with an identical free energy change, the energy difference in a single step (i.e., dimerization) dramatically alters the outcome of the polymerization reaction. As one may intuitively imagine, when $[A_0]/[B_0]$ is noticeably less than unity, unreacted B is present in much more abundance than unreacted A at equilibrium. Most impressively for the N−E mechanism, when $[B]/[A]_0$ was plotted against $([B]_0 - [A]_0)/[A]_0$, a nearly straight line with a slope close to unity was obtained (Figure 1b). This means that under the exemplified N−E conditions the amount of monomer B remaining unreacted at the polymerization equilibrium almost accounts for all B in excess of A at the beginning of the polymerization; that is, the excess B was nearly quantitatively excluded from the product molecules. It can then be inferred that repeating units A and B are present in product chains at a roughly stoichiometric ratio. In contrast, this did not occur in the isodesmic system, wherein the unreacted monomer concentration at equilibrium represents less than half of the excess monomer initially present in the system. It thus implies that the N−E system may potentially be producing polymers of higher MW than the corresponding isodesmic system under the described conditions (vide infra).

Using the equilibrium concentration data of [A], [B], and [D], the number- and weight-average degrees of polymerization as well as the polydispersity index (PDI) of the two systems were calculated (see the Supporting Information for mathematical details). In the plot of number-average degree of polymerization, $\langle dp \rangle_n$, against the total (initial) monomer ratio (Figure 2a) it was found that, even though when $[A]_0/[B]_0$ is equal to or slightly less than unity the N−E mechanism gives notably higher $\langle dp \rangle_n$ than the corresponding isodesmic system,[4] there is not much difference in $\langle dp \rangle_n$ between the two systems when $[A]_0/[B]_0$ is below 0.95. However, under the same conditions significantly higher weight-average degree of polymerization ($\langle dp \rangle_w$) was obtained for the N−E polymerization than that from the isodesmic system throughout the entire range of monomer ratio examined ($[A]_0/[B]_0$ ranging from 0.7 to 1, Figure 2b). This difference in $\langle dp \rangle_n$ and $\langle dp \rangle_w$ suggests that there is a prominent distinction in the product size distribution between the N−E and isodesmic systems. This is explicitly depicted by the polydispersity index (PDI) data (Figure 3a). While the PDI of the isodesmic polymerization remains roughly constant near 2.0 as $[A]_0/[B]_0$ diminishes from 1.0 to 0.7, the PDI of the N−E system increases to very large numbers as $[A]_0/[B]_0$ deviates from unity. The very large $\langle dp \rangle_w$ and PDI values, as well as the earlier evidence for the presence of a high concentration of unreacted excess monomer in the N−E system (Figure 1), strongly suggest that polymers with substantial MWs are generated in the N−E polymerization, coexisting with a large amount of unreacted excess monomers.

In order to confirm this hypothesis, the average degrees of polymerization were then calculated for only the product molecules (including oligomeric and polymeric species of all sizes but excluding the unreacted monomers) for the N−E and isodesmic systems. Their values are plotted against $[A]_0/[B]_0$ and shown in comparison in Figure 4 ($\langle dp \rangle_n'$ and $\langle dp \rangle_w'$ denoting the number- and weight-average degrees of polymerization of the product molecules, respectively). Just as expected, the N−E system now gives both larger number- and weight-average degrees of polymerization than the isodesmic system. Remarkably, for example when $[A]_0/[B]_0 = 0.7$, $\langle dp \rangle_n'$ equals ca. 7.5 in the isodesmic system with $\sigma = 1.0$ and $K = 1.0 \times 10^4$, while $\langle dp \rangle_n'$ reaches over 300 for the N−E mechanism with the same K value and $\sigma = 1.0 \times 10^{-4}$. Apparently, under conditions of imbalanced stoichiometry, the possible way to achieve high-MW polymers is to selectively construct molecules with comonomers at a near-stoichiometric ratio and exclude the excess monomer from the polymer chains. This is proven by

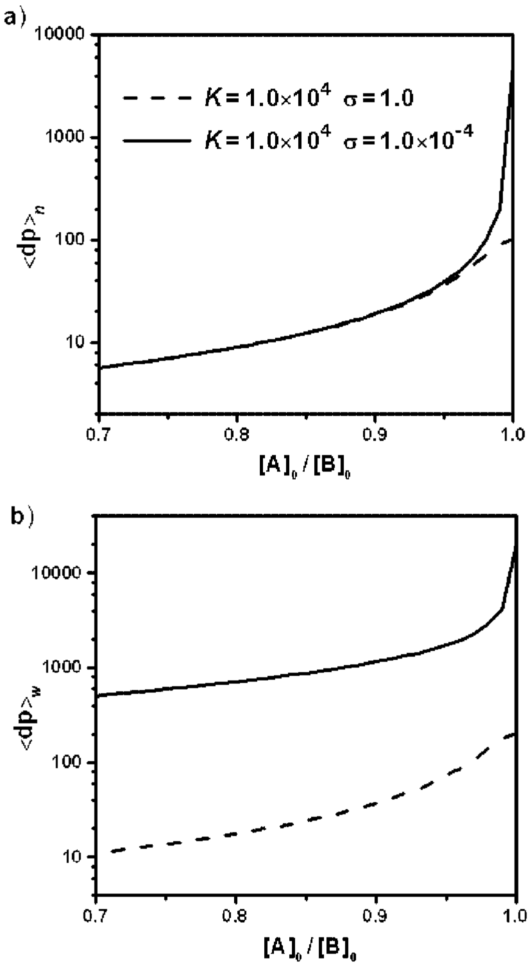


Figure 2. Plots of (a) number- and (b) weight-average degrees of polymerization (counting product molecules of all different sizes as well as unreacted monomers) as functions of initial monomer molar ratio in the nucleation−elongation (solid line) and isodesmic (dashed line) polymerizations (two lines nearly overlap in the plot of $\langle dp\rangle_n$ when $[A]_0/[B]_0 < 0.95$); representative values of equilibrium constant K and nucleation factor σ used for calculations are as indicated.

our calculations happening to the N−E polymerization. Such a monomer−polymer coexisting bimodal distribution is mechanistically achieved by creating a nucleation event in the initiation of chain growth and thereby thermodynamically destabilizing the short oligomers. Under equilibration conditions, short oligomers are thus shifted to form longer chains of higher MW by leaving out the excess terminal units.

Moreover, the PDI′ (i.e., $\langle dp\rangle_w'/\langle dp\rangle_n'$) was found to remain very close to 2.0 for the N−E polymerization as $[A]_0/[B]_0$ ranged from 1.0 to 0.7 (Figure 3b). This implies that the polymers generated from the N−E polymerization under conditions of imbalanced stoichiometry resemble the polymers produced by a classical isodesmic polymerization with a perfectly balanced functionality stoichiometry, except that the former exists in equilibrium with a larger amount of unreacted monomers. It should also be noted that Figures 1a and 4a present consistent results with those from a previously studied experimental N−E polymerization reaction of a pair of A−A and B−B type comonomers.[8] In that polycondensation reaction, a nucleation event in the chain growth resulted from the polymer product adopting a helical conformation. At polymerization equilibria with varied monomer ratios, the experimental system consistently displayed monomer−polymer bimodal distributions. Products with substantial MW were detected coexisting with considerable unreacted monomers, the amount of which was found to be nearly as much as that of the comonomer added in excess of the other at the beginning of the polymerization.

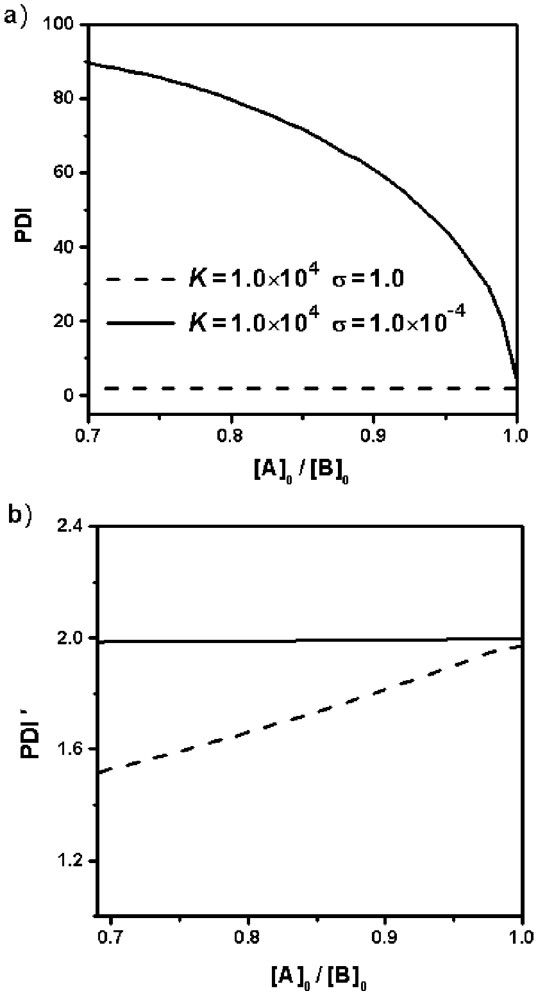


Figure 3. Plots of polydispersity index as a function of initial monomer molar ratio in the nucleation−elongation (solid line) and isodesmic (dashed line) polymerizations; representative equilibrium constant K and nucleation factor σ values used for calculations are as indicated: (a) PDI = $\langle dp\rangle_w/\langle dp\rangle_n$ and (b) PDI′ = $\langle dp\rangle_w'/\langle dp\rangle_n'$.

One may realize that the equilibrium constant for the association of two chains (e.g., $\{A_mB_n + A_pB_q \leftrightarrow A_{m+p}B_{n+q} + D\}$, with $m, n, p, q \geq 1$, $|m - n| \leq 1$, and $|p - q| \leq 1$) should be K/σ (or $4K/\sigma$, depending on the symmetry of the reactants) and argue that the observed higher MW in the N−E polymerization is merely resultant from enhanced association strength between chain ends. It is worth noting that the excess free energy gain (by a magnitude of $RT \ln \sigma$) from such a chain−chain association is the exact compensation of the extra energy paid by A_pB_q for nucleation during its formation, relative to the monomer addition as in $\{A_mB_n + A \text{ (or B)} \leftrightarrow A_{m+1}B_n \text{ (or } A_mB_{n+1}) + D\}$. Therefore, no extra energy is invested in favor of the N−E system relative to the corresponding isodesmic reaction with the same K. Nonetheless, in order to help delineate the effect of σ on polymerizations, distributions in N−E and isodesmic polymerizations with an equal K/σ but different σ value (e.g., $\sigma = 0.01$, $K = 100$, and $\sigma = 1$, $K = 10^4$) are calculated and shown in the Supporting Information (Figure S1). These results unambiguously demonstrate that even under such disfavored conditions (with a much smaller K) the N−E system still manifests larger DPs and a bimodal distribution under imbalanced stoichiometry conditions. It is therefore obvious that the observed distinctive behaviors of the N−E system are not solely resultant from an enhanced strength for chain−chain

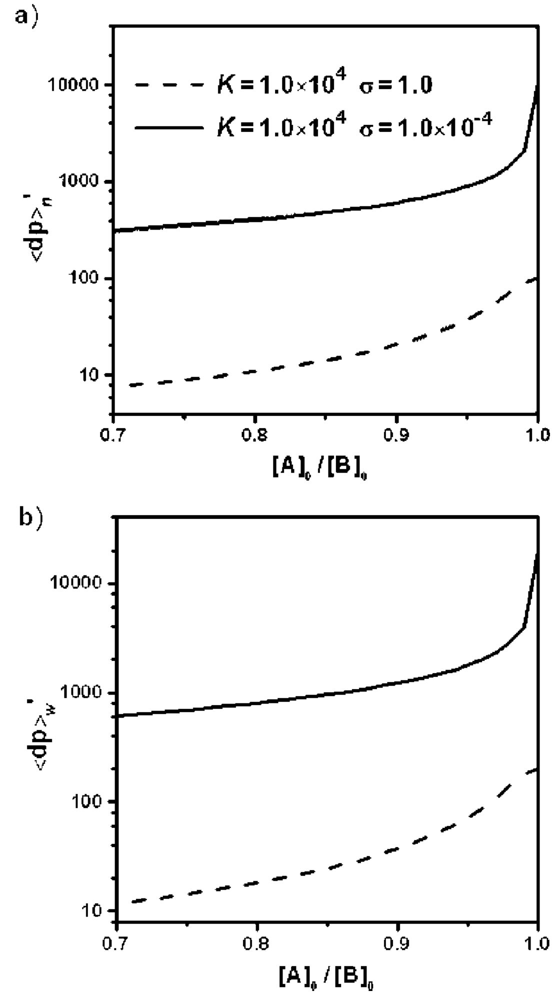


Figure 4. Plots of (a) number- and (b) weight-average degrees of polymerization of product molecules (including oligomers and polymers of different sizes but not unreacted monomers) as functions of initial monomer molar ratio in the nucleation−elongation (solid line) and isodesmic (dashed line) polymerizations; representative equilibrium constant K and nucleation factor σ values used for calculations are as indicated.

association. A significant nucleation effect, i.e., a small σ value, a large energy barrier for the initiation of chain growth, and the destabilization of the oligomeric species, is the critical prerequisite for larger DPs and a bimodal distribution under imbalanced stoichiometry conditions.

II. Statistical Approach. With the above mechanistic model, as long as the equilibrium constants of nucleation and elongation steps (i.e., σ and K) are known, all thermodynamic information about the polymerization system is obtainable. However, under experimental circumstances, σ and K are usually unknown and to be determined. What may be more directly available from the polymerization experiments are the reaction and/or functional group conversions. These may be acquired by, for example, monitoring the end-group concentration change with the assistance of various analytical techniques. Since the reaction conversion is related to the statistical probability of chain propagation, methods have been developed using statistical rules to calculate the polymer size distribution based on reaction conversion data.[10,14] Yet, most of the reaction-conversion-based models of step-growth polymerizations are built upon the principle of equal functionality reactivity.[10] That is, during a polymerization the reactivity of a functional group is independent of the size of the chain it is attached to. Consequently, in a polymerization system, regardless of their sizes, all chain ends bearing the same functionality have an equal probability to propagate (i.e., isodesmic).

This equal probability theory has proven valid in most synthetic systems.[10] Nevertheless, the essential difference of a N−E polymerization from an isodesmic system is that the reactivity of the functional groups becomes chain length dependent.[7f,15] Previously, a limited number of theoretical calculations on the product size distribution in polymerizations involving unequal functionality reactivities have been carried out.[14,15] One of the examples discussed by Case was relevant to our current problem, in which the reactivity−size dependence of acid anhydrides in polymerization with diols was analyzed.[16] In a N−E system wherein dimerization acts as the only nucleation step followed by an isodesmic chain elongation (i.e., the mechanism represented by eq 1), monomers may be considered as exhibiting similar behaviors to those of an acid anhydride in a polymerization reaction. Namely, the two functional groups are equivalent in the same monomer, but once one of them is reacted the reactivity of the other changes and remains constant thereafter.

Hence, calculations on the product distribution in such a N−E polymerization are carried out as follows. At a given overall reaction conversion, the probability of a functional group in an unreacted monomer to be reacted is different from that of the same functionality at the end of chains comprising more than one monomer units. Specifically for the N−E mechanism, the former is lower than the latter. Since we are interested in conditions of imbalanced stoichiometry, conversions of the two different functionalities, A and B, are necessarily differentiated. Consequently, four different variants related to functionality conversion, P_{A1}, P_{A2}, P_{B1}, and P_{B2}, are required for setting up a statistical model to evaluate the product size distribution in a N−E system with a nucleus size of 2. P_{A1} is designated as the fraction of monomer A that has either or both functional groups reacted; P_{A2} is the fraction of the same kind of monomer with both functional groups reacted. Thus, $(P_{A1} - P_{A2})$ is the fraction of monomer A with only one and either one of the functionalities reacted, and $(P_{A1} + P_{A2})/2$ represents the overall conversion of functionality A. P_{B1} and P_{B2} are denoted for the corresponding physical meanings related to monomer B.

In addition to P_{A1}, P_{A2}, P_{B1}, and P_{B2}, N_A and N_B are defined as the total (initial) numbers of functionalities A and B in the system, respectively (e.g., $N_A = 2[A]_0$ in a unity volume), and N_0 is the sum of N_A and N_B ($N_0 = N_A + N_B$). Since for each functionality A to be reacted, there is a functionality B being reacted, one has $N_A(P_{A1} + P_{A2}) = N_B(P_{B1} + P_{B2})$. Therefore, the total number of unreacted functionalities A and B combined, should be

$$N_P = \frac{N_A(2 - P_{A1} - P_{A2}) + N_B(2 - P_{B1} - P_{B2})}{2} \qquad (3)$$

$N_P/2$ thus corresponds to the total number of product molecules plus unreacted monomers in the system. Categorized by their end groups, there are three different series of product molecules in the system. The population of each of them can be calculated on the basis of the number of their end groups. First, the total number of end group A (i.e., unreacted functionality A) of molecules starting with a repeating unit A, ending with a repeating unit A, and containing n repeating units A and $(n - 1)$ repeating units B ($n \geq 2$) should be

$$\frac{N_0}{1 + (P_{A1} + P_{A2})/(P_{B1} + P_{B2})}\,\frac{(P_{A1} - P_{A2})}{2}\left(\frac{2P_{B2}}{P_{B1} + P_{B2}}\right)^{n-1} \times \left(\frac{2P_{A2}}{P_{A1} + P_{A2}}\right)^{n-2}\left(\frac{P_{A1} - P_{A2}}{P_{A1} + P_{A2}}\right) \qquad (4)$$

This expression is obtained as the product of the total number of functional group A, N_A (reacted and unreacted), probabilities of finding an unreacted functionality A belonging to a monomer with the other functional group reacted among all these functionality A's, finding $(n-1)$ repeating units B with both functional groups reacted among units B having at least one of the two functional groups reacted and $(n-2)$ units A with both functional groups reacted among those A having at least one functional group reacted, and last finding one unit A with only one functional group reacted among those A having at least one functional group reacted. Similarly, the total number of end group A in molecules starting with a repeating unit A and containing n repeating units A and n repeating units B $(n \geq 1)$ should be

$$\frac{N_0}{1+(P_{A1}+P_{A2})/(P_{B1}+P_{B2})}\,\frac{(P_{A1}-P_{A2})}{2}\left(\frac{2P_{B2}}{P_{B1}+P_{B2}}\right)^{n-1}\times\left(\frac{2P_{A2}}{P_{A1}+P_{A2}}\right)^{n-1}\left(\frac{P_{B1}-P_{B2}}{P_{B1}+P_{B2}}\right) \quad (5)$$

The number of end group B in molecules starting with a repeating unit B and containing n repeating units A and n repeating units B $(n \geq 1)$ is

$$\frac{N_0}{1+(P_{B1}+P_{B2})/(P_{A1}+P_{A2})}\,\frac{(P_{B1}-P_{B2})}{2}\left(\frac{2P_{A2}}{P_{A1}+P_{A2}}\right)^{n-1}\times\left(\frac{2P_{B2}}{P_{B1}+P_{B2}}\right)^{n-1}\left(\frac{P_{A1}-P_{A2}}{P_{A1}+P_{A2}}\right) \quad (6)$$

Naturally, expressions 5 and 6 are equivalent to each other as they are referring to the same set of molecules. Lastly, the number of end group B in molecules starting with a unit B and containing n repeating units A and $(n+1)$ repeating units B $(n \geq 1)$ should be

$$\frac{N_0}{1+(P_{B1}+P_{B2})/(P_{A1}+P_{A2})}\,\frac{(P_{B1}-P_{B2})}{2}\left(\frac{2P_{A2}}{P_{A1}+P_{A2}}\right)^{n}\times\left(\frac{2P_{B2}}{P_{B1}+P_{B2}}\right)^{n-1}\left(\frac{P_{B1}-P_{B2}}{P_{B1}+P_{B2}}\right) \quad (7)$$

Additionally, the total numbers of functional groups belonging to unreacted monomers A and B (neither functionality is reacted) should be respectively

$$N_A(A)=\frac{N_0(1-P_{A1})}{1+\dfrac{P_{A1}+P_{A2}}{P_{B1}+P_{B2}}} \quad (8)$$

$$N_B(B)=\frac{N_0(1-P_{B1})}{1+\dfrac{P_{B1}+P_{B2}}{P_{A1}+P_{A2}}} \quad (9)$$

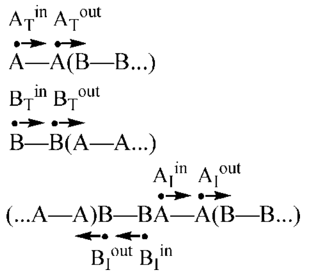
Naturally, the sum of expressions 4–9 for all n's up to infinity is equal to N_p (eq 3). Consequently, the mole and weight fractions of different molecules of any size can be calculated on the basis of the above end-group numbers along with N_p and N_0. Thus, the number- and weight-average degrees of polymerization can be derived as functions of P_{A1}, P_{A2}, P_{B1}, and P_{B2} (see the Supporting Information for details). For example, the number-average degrees of polymerization for the overall system and the product molecules (not including unreacted monomers) should be respectively

$$\langle dp\rangle_n=\frac{N_0}{N_p}=\frac{1}{1-\dfrac{(P_{A1}+P_{A2})(P_{B1}+P_{B2})}{P_{A1}+P_{A2}+P_{B1}+P_{B2}}} \quad (10)$$

$$\langle dp\rangle_n{}'=\frac{N_0-N_A(A)-N_B(B)}{N_p-N_A(A)-N_B(B)}=\frac{2P_{A1}P_{B1}+P_{A1}P_{B2}+P_{A2}P_{B1}}{P_{A1}P_{B1}-P_{A2}P_{B2}} \quad (11)$$

Here $\langle dp\rangle_n$ and $\langle dp\rangle_n{}'$ have the same definitions as in section I.

Scheme 1. Schematics Showing the Denotations for the Expected Chain Lengths after a Given Monomer Unit in a Polymerization of a Pair of Comonomers A−A and B−B[a]



[a] A−A and B−B each represents one monomer unit bearing two functionalities; units inside parentheses may, or may not, exist; subscripts T and I represent terminal (unreacted) and internal (reacted) functionalities, respectively; otherwise, the denotations are defined in accordance with previous examples of the Macosko−Miller method in the literature.[10b,17]

III. Recursive Approach (Macosko−Miller Method). Although $\langle dp\rangle_n$ and $\langle dp\rangle_n{}'$ can be obtained relatively easily, the expressions of $\langle dp\rangle_w$ and $\langle dp\rangle_w{}'$ are more sophisticated from the above method. The recursive approach developed by Macosko and Miller has proven very effective and convenient for calculating the average product size in (co)polymerizations, especially for the weight-average values.[10b,17] Herein, this recursive method is also adapted to the N−E polymerizations.

Unlike all previous examples using the Macosko and Miller method in the literature, the N−E polymerization features a size-dependent chain extension probability. Accordingly, modifications to the original method are warranted to accommodate this new characteristic. Again, in the simplest case of N−E chain growth, dimerization being the only nucleation step followed by an isodesmic elongation process, a functional group at the end of a chain that contains two monomer units or more has a reactivity (i.e., probability of being reacted and for chain extension) different than that belonging to an unreacted monomer. Namely, the likelihood of the presence of another monomer unit following a given one is dependent on the presence (or absence) of a repeating unit prior to the one of interest. Accordingly, in the modified Macosko−Miller protocol, eight different variants are necessary for adequately defining the appropriate recursive correlations (Scheme 1).

Equations showing the correlations between these variants in a N−E polymerization with a nucleus size of two can be obtained using the recursive principle:[17]

$$\begin{aligned}
A_T{}^{in}&=1+A_T{}^{out}\\
A_T{}^{out}&=\frac{P_{A1}-P_{A2}}{2-(P_{A1}+P_{A2})}B_I{}^{in}\\
A_I{}^{in}&=1+A_I{}^{out}\\
A_I{}^{out}&=\frac{2P_{A2}}{P_{A1}+P_{A2}}B_I{}^{in}\\
B_T{}^{in}&=1+B_T{}^{out}\\
B_T{}^{out}&=\frac{P_{B1}-P_{B2}}{2-(P_{B1}+P_{B2})}A_I{}^{in}\\
B_I{}^{in}&=1+B_I{}^{out}\\
B_I{}^{out}&=\frac{2P_{B2}}{P_{B1}+P_{B2}}A_I{}^{in}
\end{aligned} \quad (12)$$

In the above equations, $A/B_{T/I}{}^{in/out}$ have the definitions related to expected chain lengths as indicated in Scheme 1, and P_{A1}, P_{A2}, P_{B1}, and P_{B2} have the same meanings as designated in section II. Solving eq 12 gives solutions to all eight unknowns ($A/B_{T/I}^{in/out}$) as functions of P_{A1}, P_{A2}, P_{B1}, and P_{B2} (see the Supporting Information for details). With these results, one can calculate the weight- and number-average degrees of polymerization for the N−E system as follows:

$$\begin{aligned}\langle \mathrm{dp}\rangle_A &= \left(1-\frac{P_{A1}+P_{A2}}{2}\right)A_T^{\mathrm{in}} + \frac{P_{A1}+P_{A2}}{2}(A_I^{\mathrm{in}}+B_I^{\mathrm{in}})\\ \langle \mathrm{dp}\rangle_B &= \left(1-\frac{P_{B1}+P_{B2}}{2}\right)B_T^{\mathrm{in}} + \frac{P_{B1}+P_{B2}}{2}(A_I^{\mathrm{in}}+B_I^{\mathrm{in}})\\ \langle \mathrm{dp}\rangle_w &= \frac{[A]_0}{[A]_0+[B]_0}\langle \mathrm{dp}\rangle_A + \frac{[B]_0}{[A]_0+[B]_0}\langle \mathrm{dp}\rangle_B\end{aligned} \quad (13)$$

Here, $\langle \mathrm{dp}\rangle_A$ and $\langle \mathrm{dp}\rangle_B$ are the expected chain lengths, when a repeating unit A or B is randomly picked within the chain, respectively. In this case, unreacted monomers are included in calculations. Equations 13 and S40 accordingly give the weight- and number-average degrees of polymerization of the entire system with oligomers, polymers, and unreacted monomers all counted. If the corresponding values are to be calculated only for only the product molecules (including oligomers and polymers and excluding unreacted monomers), the following set of equations should instead be used:

$$\begin{aligned}A_T^{\mathrm{in}}(\mathrm{p}) &= 1 + A_T^{\mathrm{out}}(\mathrm{p})\\ A_T^{\mathrm{out}}(\mathrm{p}) &= B_I^{\mathrm{in}}(\mathrm{p})\\ A_I^{\mathrm{in}}(\mathrm{p}) &= 1 + A_I^{\mathrm{out}}(\mathrm{p})\\ A_I^{\mathrm{out}}(\mathrm{p}) &= \frac{2P_{A2}}{P_{A1}+P_{A2}}B_I^{\mathrm{in}}(\mathrm{p})\\ B_T^{\mathrm{in}}(\mathrm{p}) &= 1 + B_T^{\mathrm{out}}(\mathrm{p})\\ B_T^{\mathrm{out}}(\mathrm{p}) &= A_I^{\mathrm{in}}(\mathrm{p})\\ B_I^{\mathrm{in}}(\mathrm{p}) &= 1 + B_I^{\mathrm{out}}(\mathrm{p})\\ B_I^{\mathrm{out}}(\mathrm{p}) &= \frac{2P_{B2}}{P_{B1}+P_{B2}}A_I^{\mathrm{in}}(\mathrm{p})\end{aligned} \quad (14)$$

Here, the letter p in the parentheses denotes polymers. The difference between eqs 12 and 14 is that, since monomers are excluded in the latter, $A_T^{\mathrm{out}}(\mathrm{p})$ should be equal to $B_I^{\mathrm{in}}(\mathrm{p})$ and $B_T^{\mathrm{out}}(\mathrm{p})$ be equal to $A_I^{\mathrm{in}}(\mathrm{p})$. Solving eq 14, one has all the expressions for $A/B_{T/I}^{in/out}(\mathrm{p})$. Accordingly, the weight- and number-average degrees of polymerization for product molecules can be derived (see the Supporting Information).

$$\begin{aligned}\langle \mathrm{dp}\rangle_A(\mathrm{p}) &= \frac{P_{A1}-P_{A2}}{2P_{A1}}\{A_T^{\mathrm{out}}(\mathrm{p})+1\} + \frac{P_{A1}+P_{A2}}{2P_{A1}}\{A_I^{\mathrm{out}}(\mathrm{p}) + B_I^{\mathrm{out}}(\mathrm{p})+2\}\\ \langle \mathrm{dp}\rangle_B(\mathrm{p}) &= \frac{P_{B1}-P_{B2}}{2P_{B1}}\{B_T^{\mathrm{out}}(\mathrm{p})+1\} + \frac{P_{B1}+P_{B2}}{2P_{B1}}\{A_I^{\mathrm{out}}(\mathrm{p}) + B_I^{\mathrm{out}}(\mathrm{p})+2\}\\ \langle \mathrm{dp}\rangle_{w'} &= \frac{P_{A1}[A]_0}{P_{A1}[A]_0+P_{B1}[B]_0}\langle \mathrm{dp}\rangle_A(\mathrm{p}) + \frac{P_{B1}[B]_0}{P_{A1}[A]_0+P_{B1}[B]_0}\langle \mathrm{dp}\rangle_B(\mathrm{p})\end{aligned} \quad (15)$$

It should be noted that in the above analyses no conditions were set to confine P_{A1}, P_{A2}, P_{B1}, and P_{B2} to be equilibrium values. Therefore, results from the statistical and Macosko−Miller recursive methods should be applicable to pre-equilibrium states. In other words, if changes of P_{A1}, P_{A2}, P_{B1}, and P_{B2} with reaction time can be determined, the above equations are valid for evaluating the kinetic properties of polymerizations with a chain growth mechanism in accordance with eq 1. Moreover, all results derived from the last two approaches should be equally applicable to systems with or without the small-molecule byproduct formation (see the Supporting Information).

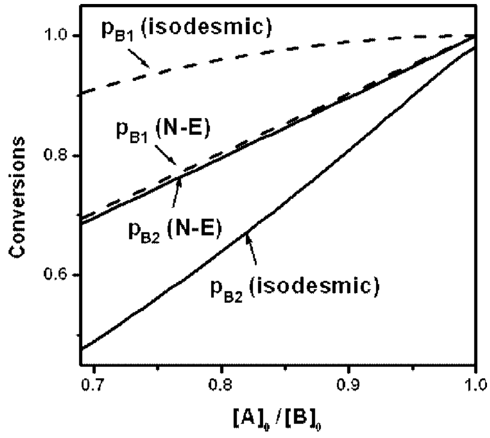


Figure 5. Equilibrium conversion data of functionality B in the nucleation−elongation ($\sigma = 1.0 \times 10^{-4}$ and $K = 1.0 \times 10^4$) and isodesmic ($\sigma = 1.0$ and $K = 1.0 \times 10^4$) polymerizations at varied initial monomer molar ratio.

As the mechanistic and statistical methods (including the Macosko−Miller recursive protocol) can all be used to describe the same type of polymerization, e.g., the N−E polymerization with a nucleus size of 2, calculation results from these different approaches should be consistent under thermodynamic equilibrium conditions. The correlations between these methods can be drawn by relating the conversion variants, P_{A1}, P_{A2}, P_{B1}, and P_{B2}, with the concentrations of A, B, and D at equilibrium. These relationships are relatively straightforward (eq S50). Combining these with eqs S3−S5, the equilibrium values of P_{A1}, P_{A2}, P_{B1}, and P_{B2} can be derived as functions of $[A]_0$, $[B]_0$, σ, and K.[11] In both isodesmic and N−E polymerizations under representative conditions as described in Figure 5, P_{A1} and P_{A2} were found to be invariably near unity at all examined $[A]_0/[B]_0$ ratios (not shown), indicating quantitative conversion achieved for the quantity-limiting monomer. Corresponding P_{B1} and P_{B2} values calculated at varied $[A]_0/[B]_0$ are shown in Figure 5. Evidently, these conversion data support the conclusions drawn from the mechanistic model. In the N−E system, there is only a very slight difference between the values of P_{B1} and P_{B2} as $[A]_0/[B]_0$ varies. Since ($P_{B1} - P_{B2}$) corresponds to the fraction of monomer B being the terminal unit of product chains and ($P_{A1} - P_{A2}$) is negligible compared to ($P_{B1} - P_{B2}$), a small value of ($P_{B1} - P_{B2}$) indicates that monomer B dominantly exists either as internal repeating units of polymers or as unreacted monomer in the N−E system and that a minimal number of chain ends and thus fairly long chains should exist in the system. In great contrast, under the isodesmic conditions P_{B2} is considerably smaller than P_{B1}, implying a much larger number of chain ends and thus a significantly lower average MW of the products. Moreover, when these conversion data were applied to the molecular size functions obtained from statistical or Macosko−Miller methods, the results (Figure S2) are in agreement with those attained from the mechanistic model shown in Figures 2 and 3.[18] The fact that independent methods give consistent results unambiguously verifies these different approaches.

Conclusions

Three models (mechanistic, statistical, and Macosko−Miller recursive methods) have been developed or modified to calculate the product size distribution in a N−E polymerization, in which dimerization acts as a nucleation event followed by an isodesmic chain elongation process. All these models consistently show that the free energy change of a single step, i.e., dimerization, dramatically alters the thermodynamic product distribution in step-growth polymerizations. The models reveal that the N−E polymerization invariably produces polymers of higher MWs than those from a classical isodesmic reaction under similar conditions. Most notably, even when the stoichiometry of polymerization functionalities is highly imbalanced, the N−E system yields polymers of substantial MWs coexisting with the excess monomer, giving rise to a bimodal distribution. This is in stark contrast to classical isodesmic systems. From the free energy point of view, a bimodal distribution is an inevitable result of nucleation. With an energy-disfavored chain initiation process, species of nucleus size are elevated in energy and destabilized relative to longer polymers and monomers. Thus, a reversible N−E polymerization favors a distribution comprising more stable, longer chains coexisting with unreacted monomers over one that mainly consists of short oligomers. This disparity in product distribution between isodesmic and N−E polymerizations is manifested most remarkably under imbalanced stoichiometric conditions as short oligomers dominate in such isodesmic reactions, and the equilibrium state is shifted away from these high-energy species in the N−E system. This feature of high tolerance of stoichiometry imbalance depicted by the N−E polymerization implies its unique potential applications in condensation polymer syntheses. The least of it may be loosening the requirement for monomer purity in step-growth polymerizations and allowing high polymers to be made under conditions of less rigorously controlled stoichiometry.

Acknowledgment. This work is partly supported by Peking University and the National Natural Science Foundation of China (No. 50603001). Part of this work was carried out at University of Illinois at Urbana−Champaign (UIUC) while DZ was a graduate student. D.Z. gratefully acknowledges Professor Jeffery S. Moore of UIUC for stimulating discussions that inspired this work. The authors also thank the anonymous reviewer for pointing out the statistical factor that should be taken into consideration in the mechanistic model, which was originally neglected.

Supporting Information Available: Detailed mathematical derivations and more plots of numerical calculation data. This material is available free of charge via the Internet at http://pubs.acs.org.

References and Notes

(1) Oosawa, F.; Asakura, S. *Thermodynamics of the Polymerization of Protein*; Academic Press: New York, 1975.

(2) (a) Klug, A. *Angew. Chem., Int. Ed. Engl.* **1983**, *22*, 565–582. (b) Caspar, D. L. *Biophys. J.* **1980**, *32*, 103–138. (c) van Workum, K.; Douglas, J. F. *Macromol. Symp.* **2005**, *227*, 1–16.

(3) For representative synthetic systems exhibiting helical supramolecular structures, see: (a) Engelkamp, H.; Middelbeek, S.; Nolte, R. J. M. *Science* **1999**, *284*, 785–788. (b) Berl, V.; Huc, I.; Khoury, R. G.; Krische, M. J.; Lehn, J. M. *Nature (London)* **2000**, *407*, 720–723. (c) Oh, K.; Jeong, K. S.; Moore, J. S. *Nature (London)* **2001**, *414*, 889–893. (d) Percec, V.; Glodde, M.; Bera, T. K.; Miura, Y.; Shiyanovskaya, I.; Singer, K. D.; Balagurusamy, V. S. K.; Heiney, P. A.; Schnell, I.; Rapp, A.; Spiess, H. W.; Hudson, S. D.; Duan, H. *Nature (London)* **2002**, *419*, 384–387. (e) Jonkheijm, P.; van der Schoot, P.; Schenning, A. P. H. J.; Meijer, E. W. *Science* **2006**, *313*, 80–83.

(4) Zhao, D.; Moore, J. S. *Org. Biomol. Chem.* **2003**, *1*, 3471–3491.

(5) For few examples of synthetic polymerizations exhibiting N−E characteristics, see: (a) Kern, W.; Jaacks, V. *J. Polym. Sci.* **1960**, *48*, 399–404. (b) Leese, L.; Baumber, M. W. *Polymer* **1965**, *6*, 269–286. (c) Iwakura, Y.; Uno, K.; Oya, M. *J. Polym. Sci., Part A1* **1967**, *5*, 2867–2874. (d) Komoto, T.; Akaishi, T.; Oya, M.; Kawai, T. *Makromol. Chem.* **1972**, *154*, 151–159. (e) de Loos, M.; van Esch, J.; Kellogg, R. M.; Feringa, B. L. *Angew. Chem., Int. Ed.* **2001**, *40*, 613–616. (f) Lortie, F.; Boileau, S.; Bouteiller, L. *Chem.−Eur. J.* **2003**, *9*, 3008–3014. (g) Simic, V.; Bouteiller, L.; Jalabert, M. *J. Am. Chem. Soc.* **2003**, *125*, 13148–13154.

(6) (a) Zhao, D.; Moore, J. S. *J. Am. Chem. Soc.* **2002**, *124*, 9996–9997. (b) Zhao, D.; Moore, J. S. *Macromolecules* **2003**, *36*, 2712–2720.

(7) For examples of previous models suitable for studying N−E polymerizations, see: (a) Oosawa, F.; Kasai, M. *J. Mol. Biol.* **1962**, *4*, 10–21. (b) Oosawa, F. *J. Theor. Biol.* **1970**, *27*, 69–86. (c) Winklmair, D. *Arch. Biochem. Biophys.* **1971**, *147*, 509–514. (d) Martin, R. B. *Chem. Rev.* **1996**, *96*, 3043–3064. (e) Edelstein-Keshet, L.; Ermentrout, G. B. *Bull. Math. Biol.* **1998**, *60*, 449–475. (f) van der Schoot, P. *Supramolecular Polymers*, 2nd ed.; Ciferri, A., Ed.; Taylor & Francis: London, 2005; Chapter 3.

(8) Zhao, D.; Moore, J. S. *J. Am. Chem. Soc.* **2003**, *125*, 16294–16299.

(9) A classical example of polymerization yielding high polymers under imbalanced stoichiometry is interfacial polymerization:(a) Morgan, P. W.; Kwolek, S. L. *J. Chem. Educ.* **1959**, *36*, 182–185. (b) Morgan, P. W.; Kwolek, S. L. *J. Polym. Sci.* **1959**, *40*, 299–327. For a more recent example of synthetic polymerization under imbalanced stoichiometry, see: (c) Kimura, K.; Kohama, S.; Yamashita, Y. *Macromolecules* **2003**, *36*, 5043–5046. (d) Kimura, K.; Kohama, S.; Yamashita, Y. *Macromolecules* **2002**, *35*, 7545–7552. In this case, crystallization of the polymer product was responsible for driving the polymer growth and excluding the monofunctional monomer.

(10) (a) Flory, P. J. *Principles of Polymer Chemistry*; Cornell University Press: Ithaca, NY, 1953. (b) Odian, G. *Principles of Polymerization*, 3rd ed.; John Wiley & Sons: New York, 1991.

(11) The calculations were conducted using Mathematica 4.2. Representative numerical solutions were obtained when the software failed to generate algebraic solutions.

(12) Here only two exemplary sets of conditions are chosen to illustrate the distinct behaviors of the two types of polymerizations; results of more diverse systems can be conveniently obtained by applying various conditions of interest to eqs S2−S13 provided in the Supporting Information.

(13) Although not obvious from Figure 1, this conclusion holds when $[A]_0/[B]_0 = 1$ (cf. ref 4).

(14) For examples, see: (a) Schaefgen, J. R.; Flory, P. J. *J. Am. Chem. Soc.* **1948**, *70*, 2709–2718. (b) Nanda, V. S.; Jain, S. C. *J. Chem. Phys.* **1968**, *49*, 1318–1320. (c) Gupta, S. K.; Kumar, A.; Bhargava, A. *Polymer* **1979**, *20*, 305–310. (d) Gupta, S. K.; Kumar, A.; Bhargava, A. *Eur. Polym. J.* **1979**, *15*, 557–564. (e) Vermonden, T.; van der Gucht, J.; de Waard, P.; Marcelis, A. T. M.; Besseling, N. A. M.; Sudhölter, E. J. R.; Fleer, G. J.; Cohen Stuart, M. A. *Macromolecules* **2003**, *36*, 7035–7044.

(15) More recently, van der Schoot et al. have developed a statistical mechanical model suitable for evaluating the helical transition of supramolecular polymers exhibiting N−E characteristics: (a) van der Schoot, P.; Michels, M. A. J.; Brunsveld, L.; Sijbesma, R. P.; Ramzi, A. *Langmuir* **2000**, *16*, 10076–10083. (b) van Gestel, J.; van der Schoot, P.; Michels, M. A. J. *J. Phys. Chem. B* **2001**, *105*, 10691–10699. (c) van Gestel, J.; van der Schoot, P.; Michels, M. A. J. *Langmuir* **2003**, *19*, 1375–1383. (d) van Jaarsveld, J.; van der Schoot, P. *Macromolecules* **2007**, *40*, 2177–2185.

(16) Case, L. C. *J. Polym. Sci.* **1958**, *29*, 455–495.

(17) (a) Macosko, C. W.; Miller, D. R. *Macromolecules* **1976**, *9*, 199–206. (b) Miller, D. R.; Macosko, C. W. *Macromolecules* **1978**, *11*, 656–662. (c) Lopez-Serrano, F.; Castro, J. M.; Macosko, C. W.; Tirrell, M. *Polymer* **1980**, *21*, 263–273. (d) Ozizmir, E.; Odian, G. *J. Polym. Sci., Polym. Chem.* **1980**, *18*, 2281–2291.

(18) When data were handled numerically, slight differences may occur between the calculation results from the three different methods due to small numerical errors.

MA071884A